MOLECULAR AND CELLULAR BIOLOGY, Mar. 1989, p. 1026–1033
0270-7306/89/031026-08$02.00/0

Vol. 9, No. 3

Isolation of Human Sequences That Replicate Autonomously in Human Cells

PATRICK J. KRYSAN, STEVEN B. HAASE, AND MICHELE P. CALOS*

Department of Genetics, Stanford University School of Medicine, Stanford, California 94305

Received 22 August 1988/Accepted 23 November 1988

We have isolated a heterogeneous collection of human genomic sequences which replicate autonomously when introduced into human cells. The novel strategy for the isolation of these sequences involved cloning random human DNA fragments into a defective Epstein-Barr virus vector. This vector alone was unable to replicate in human cells, but appeared to provide for the nuclear retention of linked DNA. The human sequences persist in a long-term replication assay (>2 months) in the presence of the viral nuclear retention sequences. Using a short-term (4-day) assay, we showed that the human sequences are able to replicate in the absence of all viral sequences. The plasmids bearing human sequences were shown to replicate based on the persistence of *Mbo*I-sensitive plasmid DNA in the long-term assay and the appearance of *Dpn*I-resistant DNA in the short-term assay. The human sequences were shown to be responsible for the replication activity and may represent authentic human origins of replication.

The structure of the chromosomal replication origins from multicellular organisms has not yet been defined. This knowledge would be basic to an understanding of the cell division involved in normal growth and development, as well as disease states such as cancer. A prerequisite to studying origin of replication structure is the molecular cloning of such sequences. In this study we describe a novel strategy that has enabled us to isolate human sequences which replicate autonomously in human cells.

The strategy of attaching genomic sequences to a selectable marker and then selecting for autonomous replication has been used successfully for the isolation of replication origins in yeasts and procaryotes (see, e.g., references 19, 40, 41, 44, and 45). This approach has not proved successful for mammalian cells (4). Selecting for an origin of replication in this manner requires not only that the DNA replicate, but also that it be maintained in the nucleus as a plasmid. It is possible that a mammalian origin, although able to mediate replication, is not able to provide for the extrachromosomal persistence of a linked plasmid. By contrast, plasmids containing the origin of replication from Epstein-Barr virus (EBV) can be readily maintained as autonomously replicating molecules in human cells (6, 7). Therefore, we hypothesized that EBV vectors may contain sequences that provide for the nuclear retention of DNA as well as replication and that we could use these viral sequences to help us isolate human replication origins.

In 1984 Yates et al. reported the isolation of a 1.7-kilobase-pair (kb) sequence, *oriP*, from the genome of EBV which acts in *cis* to permit linked DNA to replicate in human cells (46). A viral sequence encoding the Epstein-Barr nuclear antigen (EBNA-1) is required to activate replication from *oriP* (47). In human cells expressing EBNA-1, vectors containing *oriP* replicate stably as unrearranged plasmids carried at a copy number of approximately 10 per cell, as long as selection for a marker gene on the plasmid is maintained (8, 27, 36, 46, 47). EBNA-1 is the only viral protein involved in EBV replication (47), and, like the chromosomes, EBV appears to replicate once per cell cycle (1).

The structure of *oriP* consists of two components (3, 23, 27, 36). There is a 21-member family of tandem repeats, each member being a 30-base-pair sequence with an internal palindrome. This family of repeats may be involved in plasmid maintenance and can act as a transcriptional enhancer in the presence of EBNA-1 (27, 35). At a distance of 1 kb from the family of repeats lies a region of dyad symmetry consisting of a 65-bp palindrome overlapping with four partial copies of the 30-bp repeat. EBNA-1 binds to each of the repeat sequences (34). Both components of *oriP* are required in *cis* for replication (36). The dyad region is thought to be the origin of DNA synthesis by analogy to other characterized origins of replication, which contain similar palindromic structures (26, 36).

Yates and Camiolo (48) have shown that mutations in *EBNA-1* can inhibit the stable maintenance of EBV vectors without affecting their replication competence in a transient assay. These data support the idea that EBNA-1 may play a role in the nuclear retention of DNA separate from its role in replication. In other experiments, Reisman et al. (36) provided evidence consistent with the idea that a nuclear retention function may exist independently of the replication activity in EBV vectors and may be dependent on EBNA-1 and the family of repeats in *oriP*. In the presence of EBNA-1, the family of repeats alone, without the dyad, does not permit replication (27, 36). The family of repeats and *EBNA-1* do however, provide for a phenomenon known as transient drug resistance (35, 36). A population of cells transfected with a plasmid carrying the family of repeats, *EBNA-1*, and a selectable marker will survive drug selection for 2 to 3 weeks, after which only cells containing the integrated marker will survive. By contrast, for cells receiving a plasmid carrying the same marker gene without the family of repeats, only those few cells with the marker integrated show any growth throughout the selection period. This phenomenon of transient drug resistance could be due to increased expression of the marker gene as a result of the enhancer activity of the family of repeats (35, 36). Alternatively, it could be caused by prolonged retention of the transfected DNA, even in the absence of replication, mediated by the EBNA-1–*oriP* interaction.

We directly demonstrate here the ability of EBNA-1 and

* Corresponding author.

the family of repeats to retain linked DNA within human cells. This nuclear retention function interested us as a potential means of stabilizing introduced human sequences carrying replication origins. We constructed a replication-defective EBV vector by removing the dyad region of *oriP*. Random pieces of the human genome were then cloned into the vector, which still possessed the viral sequences responsible for the nuclear retention function. A large collection of human sequences was isolated that enable this vector to replicate in human cells in a long-term assay. These human sequences also replicate in the absence of all viral sequences when tested in a short-term assay.

MATERIALS AND METHODS

Plasmids. The plasmid p220.2 (8) (a gift of B. Sugden) was digested with *Eco*RV and *Hpa*I and recircularized to produce pDY$^-$. Genomic DNA from human 293S cells was partially digested with *Hin*dIII and cloned into the unique *Hin*dIII site of pDY$^-$ to make the human library, pLIB.

The 0.8-kb *Eco*RI-*Hin*dIII fragment of yeast *ARS1* from *GARS1*(U) (38) (a gift of R. W. Davis and R. J. Sapolsky) was ligated to pUCR (16), which had been digested partially with *Hin*dIII and to completion with *Eco*RI. The resulting plasmid, pUCR-ARS, was digested completely with *Hin*dIII to give a 0.8-kb fragment containing the *ARS1* sequences as well as a pUC19 polylinker, now attached to the *Eco*RI site of the yeast DNA. This 0.8-kb *Hin*dIII fragment was cloned into the unique *Hin*dIII site of pDY$^-$ to give pDY$^-$-ARS.

The 312-bp *Eco*RIIG fragment of simian virus 40 (SV40) on *Eco*RI linkers (31) was cloned into pUCR (16) which had been completely digested with *Eco*RI. The resulting plasmid, pUCR-SV40, was digested to completion with *Hin*dIII to release a 283-bp fragment containing the SV40 origin as well as a pUC19 polylinker now attached to the *Eco*RI site of the SV40 DNA. This 283-bp *Hin*dIII fragment was cloned into the unique *Hin*dIII site of pDY$^-$ to produce pDY$^-$-SV40.

The plasmid BLUR8 (37) was digested with *Bam*HI, releasing a 300-bp *Alu* fragment, which was then cloned into the unique *Bam*HI site of pDY$^-$ to produce pDY$^-$-ALU.

The plasmid pHEBo (47) was digested partially with *Eco*RV and completely with *Hpa*I and recircularized. pBODY is the resulting ligation product missing only the 140-bp EBV dyad-containing region. The plasmid p220.2 was digested with *Acc*I and recircularized to produce pORI$^-$. The plasmid pJYM (28) was digested with *Bam*HI and recircularized to produce pML.

The 14-kb *Hin*dIII human fragment from pLIB-16 was cloned into the unique *Hin*dIII sites of pBODY, pORI$^-$, and pML to produce pBODY16, pORI$^-$16, and pML16, respectively.

The p220.2, pDY$^-$, pHEBo and pORI$^-$ derivatives used in the experiment shown in Fig. 2 all carry a 5.1-kb genomic human leukocyte-associated antigen A2 (HLA-A2) *Hin*dIII fragment (24) cloned into their unique *Hin*dIII sites.

Tissue culture. The cell line 293S (39) is a suspension-adapted derivative of the human embryonic kidney cell line 293 (11). Cell line 143 is derived from a human osteosarcoma (2). All cells were grown in Dulbecco modified Eagle medium with 10% fetal calf serum, penicillin, and streptomycin in a 10% CO_2 incubator. If selection for the hygromycin resistance gene was appropriate, the media also contained hygromycin B (Calbiochem-Behring, La Jolla, Calif.) at a concentration of 200 μg/ml.

Transfections. All transfections were done by calcium phosphate coprecipitation (43) with cells which had been split 1:10 the day before. Equal molar amounts of plasmid DNA were transfected for the short-term assay, ranging from 1 to 5 μg/60-mm dish to 2.5 to 12 μg/100-mm dish. For the long-term and DNA retention assays, equal weights of plasmid DNA were transfected, 10 μg/100-mm dish or 2.5 μg/60-mm dish. The growth medium was replaced 1 day after transfection. Hygromycin selection, if done, was applied 2 to 5 days after transfection.

Determination of transfection efficiencies. Transfected cells were removed from plates with phosphate-buffered saline-EDTA (2 mM). Then 10^6 cells were stained with a monoclonal anti-HLA-A2 antibody, and then with fluorescein isothiocyanate-conjugated goat anti-mouse antibody (Sigma Chemical Co., St. Louis, Mo.). The cells were then analyzed with a fluorescence-activated cell sorter as previously described (33).

Blots and hybridization. DNA run on 0.5 to 0.7% agarose gels was transferred to GeneScreen Plus (Du Pont, NEN Research Products, Boston, Mass.) with 0.4 M NaOH as the transfer buffer. Filters were probed with plasmid DNA labeled with ^{32}P by random-primer extension (9) or nick translation (29). Hybridization took place for 12 to 18 h at 65 to 67°C in a buffer consisting of 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 1% sodium dodecyl sulfate, 5× Denhardt solution, 100 μg of denatured salmon sperm DNA per ml, and 10^7 to 10^8 cpm of probe (100 ng) (29). Washing was done under the conditions recommended by the GeneScreen manufacturer. The filters were encased in Saran Wrap and exposed to XAR-5 film (Eastman Kodak Co., Rochester, N.Y.).

RESULTS

Nuclear retention of DNA. To test the hypothesis that EBV sequences provide a nuclear retention function for plasmid DNA, we performed a time course experiment involving a direct hybridization assay for retained DNA. Four plasmids were used which could map the putative nuclear retention function to various segments of the EBV replicon. The plasmid p220.2 (Fig. 1A) contains the complete *oriP* and *EBNA-1* (8) (a gift of B. Sugden) and is fully replication competent. The plasmid pDY$^-$ is a derivative of p220.2 in which the region of dyad symmetry in *oriP* is deleted (Fig. 1B). Therefore, pDY$^-$ contains the family of repeats and *EBNA-1*, but is defective for replication. In pORI$^-$ the entire *oriP* region from p220.2 is deleted so that the plasmid contains only *EBNA-1* of EBV and cannot replicate. Finally, the plasmid pHEBo (47) contains the complete *oriP* but lacks *EBNA-1* and also does not replicate. The 5.1-kb genomic sequence encoding the human HLA-A2 cell surface antigen (24) was added to each of these plasmids. This tag allowed us to monitor the transfection efficiencies of the plasmids by using anti-HLA antibodies and fluorescence-activated cell sorter analysis (33; S. B. Haase, S. S. Heinzel, P. J. Krysan, and M. P. Calos, Mutat. Res., in press). The HLA gene also served as a common marker fragment for Southern blot hybridization.

These four plasmids were transfected into 293S cells by calcium phosphate coprecipitation. At 72 h after transfection, cells were removed from plates and samples were stained with anti-HLA antibody and assayed on the fluorescence-activated cell sorter to monitor transfection efficiency. Transfection efficiencies were similar, with an average of 54% (data not shown). This high transfection efficiency is normal for 293S cells. The remainder of the cells were washed, plated at various dilutions, and grown in the ab-

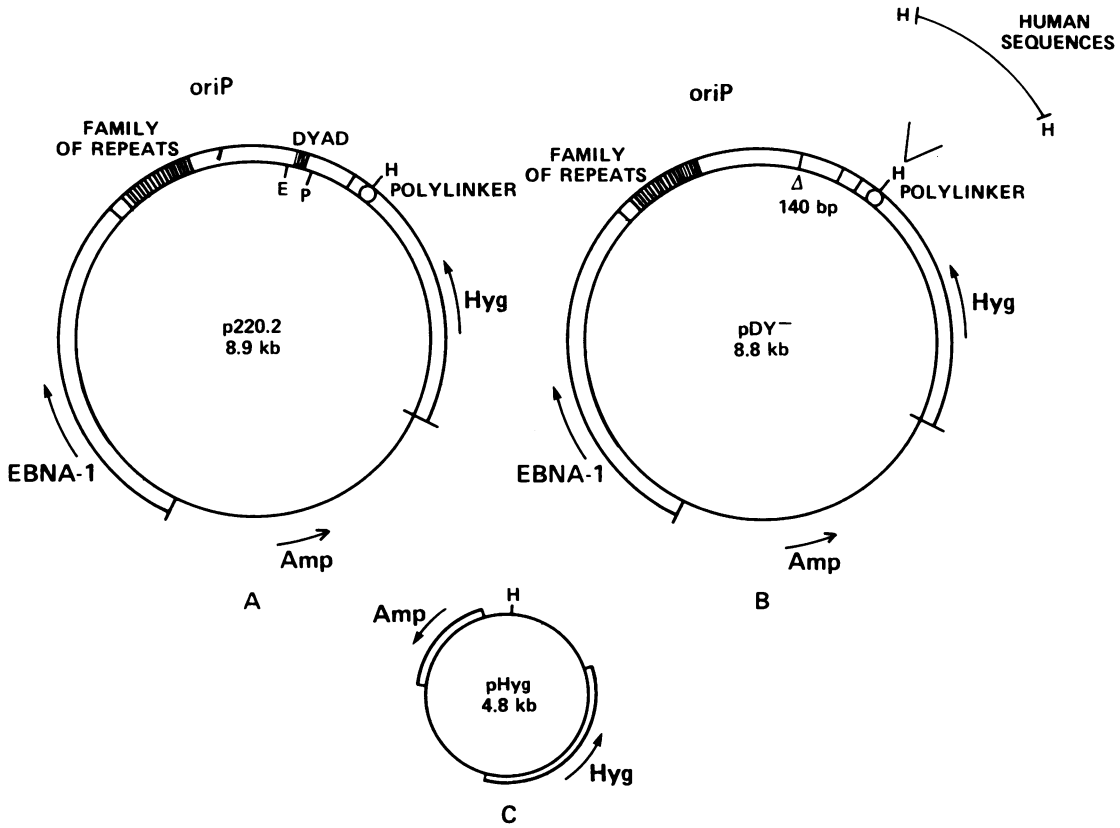

FIG. 1. Plasmid structures. (A) p220.2 has been described previously (8). It contains the *EBNA-1* gene and the complete *oriP*. The family of repeats and dyad of *oriP* are diagrammed. The dyad region is wholly contained between unique *Eco*RV and *Hpa*I sites, which are indicated. (B) pDY$^-$ was constructed by digestion of p220.2 with *Eco*RV and *Hpa*I and religation of the blunt ends. pDY$^-$ is identical to p220.2, except that it lacks the dyad symmetry region of *oriP*. The human library was constructed by ligating a *Hin*dIII partial digest of human genomic DNA to pDY$^-$ which had been cut at its unique *Hin*dIII site. (C) pHyg has been described previously (42). It carries the gene for hygromycin resistance and pBR322 sequences. No EBV sequences are present. Restriction sites: E, *Eco*RV; H, *Hin*dIII; P, *Hpa*I.

sence of selection and without splitting until harvested at later times. Plasmid DNA was isolated from cells by Hirt extraction (17) at 4, 9, 15, and 18 days after transfection. Equal fractions of Hirt extracts were digested with *Hin*dIII to release their common HLA fragment and *Mbo*I to remove replicated DNA. *Mbo*I cuts only DNA which has lost the bacterial methylation pattern. The samples were then run on an agarose gel, blotted, and probed with the 5.1-kb *Hin*dIII HLA fragment. The results are shown in Fig. 2. The bands at 5.1 kb represent unreplicated, extrachromosomal DNA remaining in the cells at the given time point. Lower-molecular-size bands represent DNA that replicated in the human cells, consequently becoming sensitive to digestion with *Mbo*I.

The majority of p220.2 DNA appeared as replicated, *Mbo*I-sensitive DNA by day 15. The appearance of *Mbo*I-sensitive bands in the pDY$^-$ samples was indicative of a low level of replication activity, well below that of p220.2. The absence of *Mbo*I-sensitive bands in the pORI$^-$ and pHEBo samples illustrated their inability to replicate. Band intensities of the 5.1-kb unreplicated band were compared by scanning densitometry. At 15 and 18 days there was approximately 10-fold more unreplicated pDY$^-$ remaining than either pORI$^-$ or pHEBo. These results show that the prolonged retention of pDY$^-$ can be attributed to the family of repeats and *EBNA-1*. That *EBNA-1* or *oriP* alone is not sufficient for nuclear retention of DNA was illustrated by the results with pORI$^-$ and pHEBo. Similar experiments with 293S cells in which *EBNA-1* is expressed in *trans* from an integrated position in the genome confirmed that it is the EBNA-1 protein and not the *EBNA-1* DNA sequence that is required for nuclear retention.

Human library construction and characterization. Our strategy for isolating human sequences capable of replication involved utilizing the unique properties of the plasmid pDY$^-$. As indicated above, the replication-defective pDY$^-$ vector appears to possess a mechanism for the prolonged nuclear retention of itself within human cells. We hoped to utilize this property of pDY$^-$ to retain autonomously replicating human sequences that in other plasmid contexts would become integrated or lost over the course of the selection period in human cells.

The starting plasmid for this work, p220.2 (Fig. 1A), replicates autonomously in human cells. It contains *EBNA-1*, the complete *oriP*, a gene conferring resistance to hygromycin, and pBR322 sequences (8). The dyad symmetry region of *oriP* was removed from p220.2 to produce pDY$^-$ (Fig. 1B). A plasmid library was then created by ligating a partial *Hin*dIII digest of human genomic DNA into the unique *Hin*dIII site of pDY$^-$. The resulting library comprised at least 10,000 independent human inserts.

The pDY$^-$ plasmid library was introduced into human 293S cells (11, 40) by calcium phosphate coprecipitation (43). As a control, pDY$^-$ was also introduced into 293S cells. The cells were placed under hygromycin selection 2 days after transfection, and hygromycin-resistant populations were

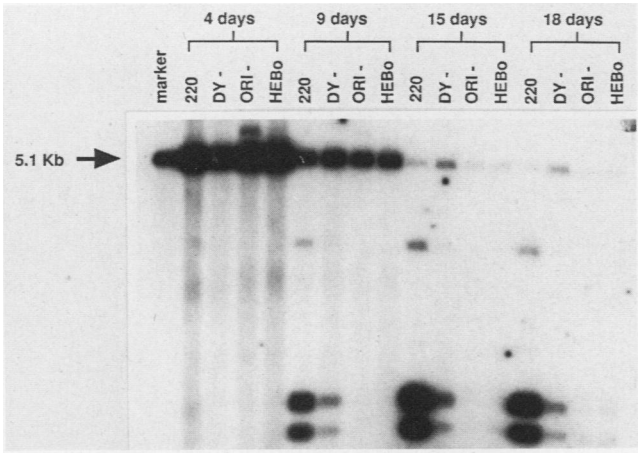

FIG. 2. Retention of DNA. The plasmids p220.2, pDY$^-$, pORI$^-$, and pHEBo were transfected into 293S cells. At 72 h later the cells were washed twice in phosphate-buffered saline–EDTA and twice in warmed medium to remove residual calcium phosphate and DNA and then plated for harvest at later time points. 250K, 200K, 100K, and 50K cells were seeded on 60-mm dishes and grown without selection for harvest at 4, 9, 15, and 18 days, respectively. Low-molecular-weight DNA was isolated by Hirt extraction (17). One-third of each Hirt extract was digested with *Mbo*I and *Hin*dIII and run on a 0.7% agarose gel. The gel was transferred to a nylon membrane and probed with a ^{32}P-labeled 5.1-kb *Hin*dIII HLA fragment. The marker lane contains 0.1 ng of *Hin*dIII-digested, HLA-tagged p220.2.

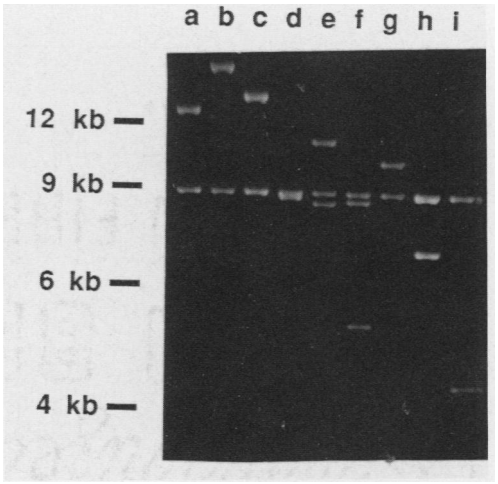

FIG. 3. Characterization of human inserts. A set of pLIB plasmids (lanes a to i) was digested with *Hin*dIII and run on a 0.5% agarose gel. The plasmids share a common 8.9-kb pDY$^-$ vector fragment. All the other bands represent human DNA. Lanes: a, pLIB9; b, pLIB13; c, pLIB16; d, pLIB36; e, pLIB41; f, pLIB75; g, pLIB6; h, pLIB21; i, pLIB136.

grown. The populations were split 1:5 approximately once per week. Following 2 months of passage, plasmid DNA was extracted from the cells by the method of Hirt (17). These extracts were then treated with *Dpn*I and used to transform bacteria to determine whether any free plasmid DNA remained in the cells. The enzyme *Dpn*I cuts only DNA which retains a bacterial modification pattern, which is lost when DNA replicates in a human cell. The *Dpn*I digestion thereby ensured that unreplicated input DNA would not be rescued. Only one bacterial colony was recovered from the pDY$^-$ population, whereas 77 colonies were recovered from the population which received pDY$^-$ containing human sequences. This result suggested that pDY$^-$ underwent little or no replication in the human cells but that some of the pDY$^-$ plasmids containing human genomic inserts were replicating. The plasmids which had replicated in human cells were called pLIBs, with each human insert being given a number. The replication behavior of individual pLIB clones is shown below.

A collection was made of bacterial colonies that had been transformed by *Dpn*I-resistant plasmid DNA recovered from the human cell population carrying the human library. To examine the structure of the human inserts, plasmid DNA was isolated from each of 400 of these colonies and digested with *Hin*dIII to release the human sequences. The *Hin*dIII digests for some of the plasmids are shown in Fig. 3. The average total insert size was 12.3 kb. The average insert size in the starting library was 6.3 kb. The collection of plasmids that emerged from the replication selection was highly heterogeneous. There appeared to be a large number of human fragments that could provide for autonomous replication of the pDY$^-$ vector.

Long-term replication. The replication behavior of several of the individual pLIB clones was examined in more detail. DNA from each of five pLIB plasmids, as well as pDY$^-$ and p220.2 controls, was reintroduced into 293S cells, and hygromycin selection was carried out. Hygromycin-resistant cells were harvested 1 and 2 months after transfection. Plasmid DNA was Hirt extracted, run on an agarose gel undigested or digested with *Mbo*I, blotted to a filter, and probed with ^{32}P-labeled p220.2 DNA. *Mbo*I cuts only DNA which has lost its bacterial methylation pattern. Therefore, sensitivity to this enzyme is diagnostic of replication in human cells. Each of the pLIB plasmids demonstrated replication (Fig. 4). However, at this 2-month time point no pDY$^-$ remained. In the absence of replication, the transfected pDY$^-$ DNA was presumably diluted to below the level of detection of this experiment. The undigested lanes show that p220.2 and the pLIB plasmids existed as extrachromosomal circles, while the *Mbo*I-digested lanes show that all of the DNA detectable at this time has replicated in the human cells and is not residual input DNA. The pLIB plasmids were variable in the levels of plasmid DNA maintained.

Cell line 293S was used for these studies because it transfects well and expresses introduced DNA efficiently. It is possible that the adenovirus *E1A* gene used to immortalize 293S cells makes these cells more permissive for replication (25). Therefore, we tested the replication of p220.2, pDY$^-$, pLIB9, and pLIB16 in human cell line 143, an osteosarcoma cell line not immortalized with *E1A* (2). These experiments showed that vector DNA was present several weeks after transfection and hygromycin selection for p220.2, pLIB9, and pLIB16, but not pDY$^-$. This DNA was fully *Mbo*I sensitive, indicating replication. This result shows that replication of the pLIB human library is not confined to 293S cells and occurs in two widely divergent human cell types.

Short-term replication. We next asked whether the human inserts could mediate the replication of linked DNA in the absence of any EBV sequences. The *Hin*dIII human insert from pLIB16 was moved to the *Hin*dIII site of pML, a pBR322 derivative lacking EBV sequences (28), producing pML16. Replication of pML, pML16, pLIB16, and p220.2 in 293S cells was compared in a short-term replication assay (Fig. 5). Plasmid DNA was extracted from the human cells 4

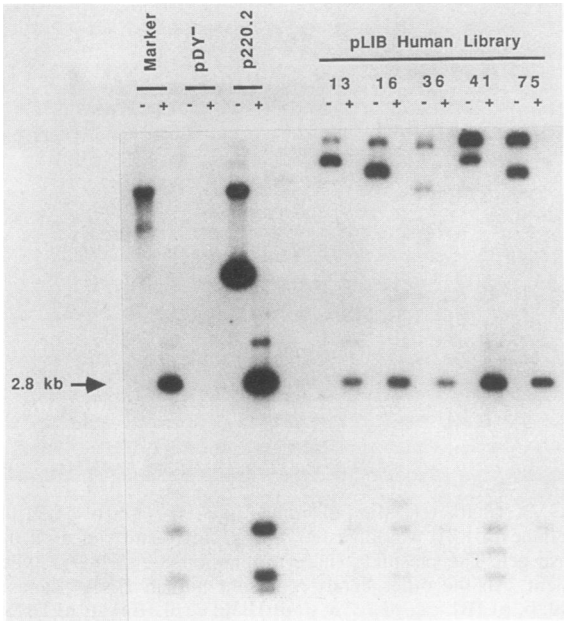

FIG. 4. Long-term replication of pLIB plasmids. 293S cells were transfected with the designated plasmids, placed under hygromycin selection, and passaged for one (36, 41, 75) or two (pDY⁻, p220.2, 13, 16) months. The Hirt extract from one-quarter of a 60-mm dish of cells was run uncut (lanes −) or cut with *Mbo*I (lanes +) on a 0.7% agarose gel, blotted, and probed with ^{32}P-labeled p220.2 DNA. The marker lanes (M) contain 1 ng of uncut (lanes −) and *Dpn*I-cut (lanes +) p220.2 DNA. No pDY⁻ DNA was detectable after 2 months, whereas replicated DNA was present for p220.2 and all the pLIB plasmids. The arrow points to the largest (2.8-kb) *Mbo*I band, indicative of replication.

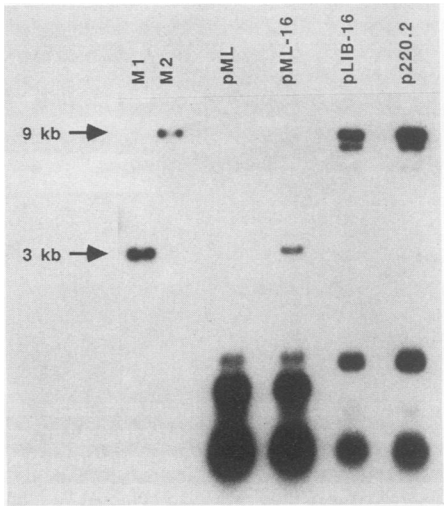

FIG. 5. Transient replication of human insert 16. 293S cells were transfected with the designated plasmids and grown for 4 days. One-third of the Hirt extract from a 60-mm dish of cells was run in each lane of the gel, blotted, and probed as described in the text. The markers are 0.5 ng of pML and 2.5 ng of p220.2 linearized with *Hin*dIII (lanes M1 and M2, respectively). The test plasmids were digested with *Hin*dIII to release their human inserts and/or linearize them and with *Dpn*I to remove unreplicated input DNA. The hybridization probe was pML. A substantial amount of *Dpn*I-resistant DNA is present for pML16, pLIB16, and p220.2, indicating replication. All of the pML DNA is sensitive to *Dpn*I, indicating lack of replication. The positions of the pML (3.0 kb) and p220.2 (8.9 kb) vector bands are indicated.

days after transfection, digested with *Hin*dIII and *Dpn*I, run on an agarose gel, blotted, and probed with ^{32}P-labeled pML DNA. The *Hin*dIII treatment linearized the vectors and released the human inserts, while *Dpn*I removed unreplicated input DNA. All of the pML DNA was sensitive to *Dpn*I digestion, indicating that it had not replicated in the human cells. Replication of pML16, pLIB16, and p220.2 was, however, easily detectable in the form of *Dpn*I-resistant vector DNA. Similar results were obtained when human insert 9 was placed into pHyg (Fig. 1C), a vector with no EBV sequences. Vector pHyg showed little replication on its own but a significant level of replication when linked to human insert 9 (data not shown). These results argue against the interpretation that we have cloned sequences similar to the EBV *oriP* that interact with EBNA-1 to mediate replication. Instead, the human inserts we have isolated mediate replication on their own, in the absence of all EBV sequences. They are autonomously replicating.

Control plasmids. To provide a more complete understanding of how the pDY⁻ vector allows replicating human sequences to be stably maintained over time, we constructed a number of control plasmids. These controls addressed two main questions: (i) what components of the pDY⁻ vector are required to allow long-term replication of linked human sequences, and (ii) what types of insert sequences allow the DY⁻ vector to replicate.

The two viral components of the pDY⁻ vector are the EBV family of repeats and *EBNA-1*. Three new vectors were enlisted to determine which, if any, of these viral sequences were needed to allow long-term replication of a linked human sequence. The plasmid pBODY is essentially pDY⁻ without *EBNA-1*. pORI⁻ is pDY⁻ without the EBV family of repeats, and pHyg (Fig. 1C) is a plasmid with no EBV sequences at all. The human insert from pLIB16 was moved into each of these vectors, giving pBODY16, pORI⁻16, and pHyg16. These plasmids were tested for short- and long-term replication ability as described above. All three vectors carrying human insert 16 were unable to persist as autonomously replicating plasmids in the long-term assay (Table 1). We have shown, however, that human insert 16 does replicate long term in the pDY⁻ context. These results indicated that both the EBV family of repeats and *EBNA-1* are required to allow autonomously replicating human sequences to be stably maintained as extrachromosomal molecules over time. Although none of these control plasmids supported long-term replication, in all cases the plasmid linked to human insert 16 demonstrated replication in the short-term assay, verifying that the replication activity is due to the human sequence. It should be noted that the same sequences found to be required for long-term replication, the family of repeats and *EBNA-1*, were also those shown above to be required for the nuclear retention of DNA.

The second group of controls focused on the pDY⁻ vector. We attempted to substitute the replication activity of the human inserts in pDY⁻ with previously defined sequences potentially involved with human replication. Yeast *ARS1*, a copy of the human *Alu* repeat, and the SV40 origin of replication were all cloned into pDY⁻ and tested for long-term replication as above. Yeast *ARS1* and human *Alu* were tested in both orientations. Yeast *ARS1* is one of the best-characterized autonomously replicating sequences isolated from *Saccharomyces cerevisiae* (38, 44). If origins of replication were completely conserved among eucaryotes,

TABLE 1. Replication assay

Plasmid	EBV sequences[a]			Human insert[b]	Origin substitute	Replication[c]	
	EBNA	*oriP*					
		Family	Dyad			Short-term	Long-term
p220.2	√	√	√			+	+
pDY⁻	√	√				–[d]	–
pLIB16	√	√		√		+	+
pORI⁻	√					–	–
pORI⁻16	√			√		+	–
pBODY		√				–	–
pBODY16		√		√		+	–
pHyg						–	–
pHyg16				√		+	–
pDY⁻-ARS	√	√			Yeast *ARS1*	–	–
pDY⁻-ALU	√	√			Human *Alu*	–	–
pDY⁻-SV40	√	√			SV40 origin	–	–

[a] Check marks indicate which EBV sequences are present on each plasmid. EBNA, Gene coding for EBNA-1; family, family of repeats from *oriP*; dyad, dyad region from *oriP*.

[b] Human insert refers to the autonomously replicating human sequence from pLIB16.

[c] +, given plasmid replicated in the specified assay (short- and/or long-term); –, the plasmid did not replicate.

[d] pDY⁻ shows a low level of background replication in transient assays.

they might be expected to replicate in our assay in human cells. The 300-bp *Alu* sequences which are highly repeated and dispersed within the genome have often been suggested to function as origins of replication (22). We wanted to test whether an *Alu* sequence could activate the replication of pDY⁻. The SV40 origin was tested because it contains a palindrome (26), and we wanted to learn whether replacing the dyad of *oriP* with another sequence that could assume a dyad conformation could restore replication ability to pDY⁻. None of these sequences allowed the linked pDY⁻ vector to replicate in the long-term assay (Table 1). From this experiment we concluded that the human inserts we isolated by using the pDY⁻ vector are either more complex than or completely different from the yeast *ARS1*, human *Alu*, and SV40 origin sequences tested.

DISCUSSION

The strategy of utilizing a plasmid carrying a defective viral origin of replication was successful in isolating components of human chromosomes which replicate autonomously in human cells. The EBV family of repeats and EBNA-1 appear to be required for the long-term establishment of the human sequences isolated in this study (Table 1). The time course experiment in Fig. 2 suggests that EBNA-1 and *oriP* enable linked DNA to be retained as plasmids for a prolonged period intracellularly, presumably in the nucleus, even if the DNA is unable to replicate. We hypothesize that via specific binding of EBNA-1 to its binding sites in *oriP* and nonspecific binding to chromatin, EBNA-1 can mediate the association of *oriP*-bearing plasmids with the chromosomes. It is known that EBNA-1 binds tightly to the family of repeats and has nonspecific DNA-binding activity (34). Furthermore, EBNA-1 can form complexes which have been demonstrated to cross-link separate pieces of DNA containing EBNA-1-binding sites (30). Alternatively, tight binding of *oriP* to the nuclear protein EBNA-1 might itself assist in nuclear localization of linked sequences. In either case, the net result would be a tendency for plasmids bearing EBNA-1 and its binding sites to be retained in the nucleus, whereas plasmids lacking such signals would be lost, either during the breakdown of the nuclear membrane at mitosis or through nuclear pores during interphase.

Such a mechanism for nuclear retention of DNA contrasts with that used by replication origins located on native chromosomes. In this case, large chromosome size and linkage to a centromere ensure that loss will be rare. EBV itself may use EBNA-1 and *oriP* as elements in its stable retention. It has been observed that EBNA-1 is found associated with metaphase chromosomes (12) and that EBV genomes are tightly but noncovalently associated with essentially all human chromosomes (13).

The presence of a putative EBV DNA retention function would explain the transient resistance phenomenon observed by Reisman et al. (36). It also appears to be responsible for our ability to isolate human sequences in a functional assay for replication. The absence of a mechanism for DNA retention probably explains why these sequences did not emerge in previous attempts (see, e.g., reference 4) to clone origins of replication from mammalian cells. The inability to clone these sequences cannot be accounted for by the rarity of origins. Fiber autoradiography studies indicate that mammalian origins are likely to be spaced 15 to 500 kb apart (20). Since the human genome contains 6.6×10^6 kb of DNA, at least 10^4 origins are expected. Our studies show that even vectors carrying sequences capable of replication will not be scored as autonomously replicating plasmids in a long-term assay in the absence of a nuclear retention function. Instead, such plasmids will either become integrated in the genome or be lost. Although they are highly unstable, *ARS* plasmids in *S. cerevisiae* are maintained sufficiently to be scored as autonomously replicating plasmids, perhaps because of the lack of nuclear membrane breakdown at mitosis in *S. cerevisiae*.

The need for a nuclear retention function to score replication in a long-term (1- to 2-month) assay left open the possibility that our human sequences alone were insufficient to mediate replication. However, experiments involving a short-term assay showed that the human inserts we isolated mediate replication even in the context of vectors lacking all viral sequences. These experiments show that the human inserts are necessary and sufficient for autonomous replication.

That our library of human inserts with replication activity is large and heterogeneous is expected on the basis of the large number of origins present in the human genome. Origins of replication in *Escherichia coli* and *S. cerevisiae* involve several hundred base pairs or less (44, 45). Our rescue of fragments averaging 12 kb from a starting input library with inserts of 6 kb may mean that more than several hundred base pairs are involved in origin function in mammalian cells. Large origins may reflect the need for multiple signals dispersed over a large area (38). Such signals may be directly involved in replication or they may be involved in other functions that affect replication.

We obtained negative results when the yeast *ARS1* sequence and a human *Alu* repeat were tested for replication in our assay. This result argues that these sequences are not recognized as origins of replication in human cells. However, this conclusion must be qualified by the proviso that the small size of the fragments tested (300 and 800 bp, respectively) or some other limitation of the assay may have prevented a positive result.

Our results should be compared with other efforts to

isolate chromosomal origins of replication from mammalian cells. These attempts involved either the isolation of one origin made accessible by fortuitous circumstances (7, 15) or the development of assays that would isolate a large set of origin fragments (18, 49). In the former category are attempts to localize the origin of replication thought to be in the amplification unit containing the gene for dihydrofolate reductase (*DHFR*). Time course experiments measuring the incorporation of radioactive nucleotides into restriction fragments have localized a putative origin to a 4.3-kb fragment located 14 kb downstream of *DHFR* in Chinese hamster ovary cells containing amplified *DHFR* genes (6, 15). However, replication of this fragment when reintroduced into CHO cells has not been demonstrated.

In the absence of a genetic assay for origin isolation, biochemical assays were developed with the goal of isolation of nascent DNA fragments, which should include replication origins (49). Some of the monkey sequences isolated by such methods appear to show replication activity above background when reintroduced into mammalian cells in a short-term assay (10). However, long-term replication has not been demonstrated.

Unlike previous studies, the human sequences isolated in our work clearly mediate both transient replication when reintroduced into human cells and long-term replication when linked to *EBNA-1* and the family of repeats from *oriP*. We see a relatively stable copy number of approximately 1 to 50 per cell, similar to or less than the copy number observed for wild-type EBV vectors and consistent with replication once (or less) per cell cycle.

We have developed a novel strategy for the isolation of human sequences which replicate autonomously in human cells. Future studies will investigate the nature of these sequences. Density-labeling experiments will address whether our sequences replicate once per cell cycle and whether they retain a specific time of replication within the DNA synthesis period of the cell cycle, as has been demonstrated for many chromosomal replicons (14). A novel two-dimensional gel technique (5, 21, 32) may permit us to map the point where replication initiates to within our human sequences. In addition, subcloning and sequence analysis of the human inserts, as well as hybridization studies between them and the chromosomes, should help clarify the relationship, if any, between the sequences we have isolated and chromosomal origins of replication.

ACKNOWLEDGMENTS

We thank Lun Lam and Charles Andre for contributing data, Scott Heinzel and Bob DuBridge for helpful discussions, and Jo Ann Katheiser for secretarial assistance.

This work was supported by Public Health Service grants (CA 33056 and CA 44645) from the National Cancer Institute to M.P.C.

LITERATURE CITED

1. **Adams, A.** 1987. Replication of latent Epstein-Barr virus genomes in Raji cells. J. Virol. **61**:1743–1746.
2. **Bacchetti, S., and F. L. Graham.** 1977. Transfer of the gene for thymidine kinase to thymidine kinase-deficient human cells by purified herpes simplex viral DNA. Proc. Natl. Acad. Sci. USA **74**:1590–1594.
3. **Baer, R., A. T. Bankier, M. D. Biggin, P. L. Deininger, P. J. Farrell, T. J. Gibson, G. Hatful, G. S. Hudson, S. C. Satchwell, C. Seguin, P. S. Tuffnell, and B. G. Barrell.** 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature (London) **310**:207–211.
4. **Biamonti, G., G. Della Valle, D. Talarico, F. Cobianchi, S. Riva, and A. Falaschi.** 1985. Fate of exogenous recombinant plasmids introduced into mouse and human cells. Nucleic Acids Res. **13**:5545–5561.
5. **Brewer, B. J., and W. L. Fangman.** 1987. The localization of replication origins on ARS plasmids in *S. cerevisiae*. Cell **51**:463–471.
6. **Burhans, W. C., J. E. Selegue, and N. H. Heintz.** 1986. Isolation of the origin of replication associated with the amplified Chinese hamster dihydrofolate reductase domain. Proc. Natl. Acad. Sci. USA **83**:7790–7794.
7. **Carroll, S. M., P. Gaudray, M. L. De Rosa, J. F. Emery, J. L. Meinkoth, E. Nakkim, M. Subler, D. D. Von Hoff, and G. M. Wahl.** 1987. Characterization of an episome produced in hamster cells that amplify a transfected CAD gene at high frequency: functional evidence for a mammalian replication origin. Mol. Cell. Biol. **7**:1740–1750.
8. **DuBridge, R. B., P. Tang, H. C. Hsia, P.-M. Leong, J. H. Miller, and M. P. Calos.** 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle vector. Mol. Cell. Biol. **7**:379–387.
9. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
10. **Frappier, L., and M. Zannis-Hadjopoulos.** 1987. Autonomous replication of plasmids bearing monkey DNA origin-enriched sequences. Proc. Natl. Acad. Sci. USA **84**:6668–6672.
11. **Graham, F. L., J. Smiley, W. C. Russell, and R. Nairn.** 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J. Gen. Virol. **36**:59–72.
12. **Grogan, E. A., W. P. Summers, S. Dowling, D. Shedd, L. Gradoville, and G. Miller.** 1983. Two Epstein-Barr virus nuclear neo-antigens distinguished by gene transfer, serology and chromosome binding. Proc. Natl. Acad. Sci. USA **80**:7650–7653.
13. **Harris, A., B. D. Young, and B. E. Griffin.** 1985. Random association of Epstein-Barr virus genomes with host cell metaphase chromosomes in Burkitt's lymphoma-derived cell lines. J. Virol. **56**:328–332.
14. **Hatton, K. S., V. Dhar, E. H. Brown, M. A. Iqbal, S. Stuart, V. T. Didamo, and C. L. Schildkraut.** 1988. Replication program of active and inactive multigene families in mammalian cells. Mol. Cell. Biol. **8**:2149–2158.
15. **Heintz, N. H., and J. L. Hamlin.** 1982. An amplified chromosomal sequence that includes the gene for dihydrofolate reductase initiates replication within specific restriction fragments. Proc. Natl. Acad. Sci. USA **79**:4083–4087.
16. **Heinzel, S. S., P. J. Krysan, M. P. Calos, and R. B. DuBridge.** 1988. Use of simian virus 40 replication to amplify Epstein-Barr virus shuttle vectors in human cells. J. Virol **62**:3738–3746.
17. **Hirt, B.** 1967. Selective extraction of polyma DNA from infected mouse cultures. J. Mol. Biol. **26**:365–369.
18. **Holst, A., F. Muller, G. Zastrow, H. Zentgraf, S. Schwender, E. Dinki, and F. Grummt.** 1988. Murine genomic DNA sequences replicating autonomously in mouse L cells. Cell **52**:355–365.
19. **Hsiao, C., and J. Carbon.** 1979. High frequency transformation of yeast by plasmids containing the cloned yeast *ARG4* gene. Proc. Natl. Acad. Sci. USA **76**:3829–3833.
20. **Huberman, J. A., and A. D. Riggs.** 1968. On the mechanism of DNA replication in mammalian chromosomes. J. Mol. Biol. **32**:327–341.
21. **Huberman, J. A., L. D. Spotila, K. A. Nawotka, S. M. El-Assouli, and L. R. Davis.** 1987. The *in vivo* replication origin of yeast 2μm plasmid. Cell **51**:473–481.
22. **Jelinek, W. R., T. P. Toomey, L. Leinwand, C. H. Duncan, P. A. Biro, P. V. Choudary, S. M. Weissmann, C. M. Rubin, C. M. Houck, P. L. Deininger, and C. W. Schmid.** 1980. Ubiquitous, interspersed repeated sequences in mammalian genomes. Proc. Natl. Acad. Sci. USA **77**:1398–1402.
23. **Karlin, S., and B. E. Blaisdell.** 1987. A model for the development of the tandem repeat units in the EBV *oriP* region and a discussion of their possible function. J. Mol. Evol. **25**:215–229.
24. **Koller, B. H., and H. T. Orr.** 1985. Cloning and complete sequence of an HLA-A2 gene: analysis of two HLA-A alleles at the nucleotide level. J. Immunol. **134**:2727–2733.
25. **Lebkowski, J. S., S. Clancy, and M. P. Calos.** 1985. SV40

replication in adenovirus-transformed human cells antagonizes gene expression. Nature (London) **317**:169–171.
26. **Li, J. J., K. W. C. Peden, R. A. F. Dixon, and T. Kelly.** 1986. Functional organization of the simian virus 40 origin of DNA replication. Mol. Cell. Biol. **6**:1117–1128.
27. **Lupton, S., and A. J. Levine.** 1985. Mapping genetic elements of Epstein-Barr virus that facilitate extrachromosomal persistence of Epstein-Barr virus-derived plasmids in human cells. Mol. Cell. Biol. **5**:2533–2542.
28. **Lusky, M., and M. Botchan.** 1981. Inhibition of SV40 replication in simian cells by specific pBR322 DNA sequences. Nature (London) **293**:79–81.
29. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Milman, G., and E. S. Hwang.** 1987. Epstein-Barr virus nuclear antigen forms a complex that binds with high concentration dependence to a single DNA-binding site. J. Virol. **61**:465–471.
31. **Myers, R. M., and R. Tjian.** 1980. Construction and analysis of simian virus 40 origins defective in tumor antigen binding and DNA replication. Proc. Natl. Acad. Sci. USA **77**:6491–6495.
32. **Nawotka, K. A., and J. A. Huberman.** 1988. Two-dimensional gel electrophoretic method for mapping DNA replicons. Mol. Cell. Biol. **8**:1408–1413.
33. **Parks, D. R., L. L. Lanier, and L. A. Herzenberg.** 1986. Flow cytometry and fluorescence activated cell sorting (FACS), p. 29.1–29.21. *In* D. M. Weir, L. A. Herzenberg, C. C. Blackwell, and L. A. Herzenberg (ed.), The handbook of experimental immunology, 4th ed. Blackwell Scientific Publications, Ltd., Edinburgh.
34. **Rawlins, D. R., G. Milman, S. D. Hayward, and G. S. Hayward.** 1985. Sequence-specific DNA binding of the Epstein-Barr nuclear antigen (EBNA-1) to clustered sites in the plasmid maintenance region. Cell **42**:859–868.
35. **Reisman, D., and B. Sugden.** 1986. *trans*-Activation of an Epstein-Barr viral transcriptional enhancer by the Epstein-Barr viral nuclear antigen 1. Mol. Cell. Biol. **6**:3838–3846.
36. **Reisman, D., J. Yates, and B. Sugden.** 1985. A putative origin of replication of plasmids derived from Epstein-Barr is composed of two *cis*-acting components. Mol. Cell. Biol. **5**:1822–1832.
37. **Rubin, C. M., C. M. Houck, P. L. Deininger, T. Friedman, and C. W. Schmid.** 1980. Partial nucleotide sequence of the 300-nucleotide interspersed repeated human DNA sequences. Nature (London) **284**:372–374.
38. **Snyder, M., R. J. Sapolsky, and R. W. Davis.** 1988. Transcription interferes with elements important for chromosome maintenance in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8**:2184–2194.
39. **Stillman, B. W., and Y. Gluzman.** 1985. Replication and supercoiling of simian virus 40 DNA in cell extracts from human cells. Mol. Cell. Biol. **5**:2051–2060.
40. **Stinchcomb, D. T., K. Struhl, and R. W. Davis.** 1979. Isolation and characterization of a yeast chromosome replicator. Nature (London) **282**:39–43.
41. **Stinchcomb, D. T., M. Thomas, J. Kelly, E. Selker, and R. W. Davis.** 1980. Eukaryotic DNA segments capable of autonomous replication in yeast. Proc. Natl. Acad. Sci. USA **77**:4559–4563.
42. **Sugden, B., K. Marsh, and J. Yates.** 1985. A vector that replicates as a plasmid and can be efficiently selected in B lymphoblasts transformed by Epstein-Barr virus. Mol. Cell. Biol. **5**:410–413.
43. **Wigler, M., R. Sweet, G. K. Sim, B. Wold, A. Pellicer, E. Lacy, T. Maniatis, S. Silverstein, and R. Axel.** 1979. Transformation of mammalian cells with genes from prokaryotes and eukaryotes. Cell **16**:777–785.
44. **Williamson, D. H.** 1985. The yeast ARS element, six years on: a progress report. Yeast **1**:1–14.
45. **Yasuda, S., and Y. Hirota.** 1977. Cloning and mapping of the replication origin of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **74**:5458–5462.
46. **Yates, J., N. Warren, D. Reisman, and B. Sugden.** 1984. A *cis*-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. Proc. Natl. Acad. Sci. USA **81**:3806–3810.
47. **Yates, J. L., N. Warren, and B. Sugden.** 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. Nature (London) **313**:812–815.
48. **Yates, J. L., and S. M. Camiolo.** 1988. Dissection of DNA replication and enhancer activation functions of Epstein-Barr virus nuclear antigen 1. Cancer Cells **6**:197–205.
49. **Zannis-Hadjopoulos, M., M. Persico, and R. G. Martin.** 1981. The remarkable instability of replication loops provides a general method for the isolation of origins of DNA replication. Cell **27**:156–163.